

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-183886  
 (43)Date of publication of application : 30.06.2000

(51)Int. CI.

H04L 12/28

(21)Application number : 10-360232

(71)Applicant : SUMITOMO ELECTRIC IND LTD

(22)Date of filing : 18.12.1998

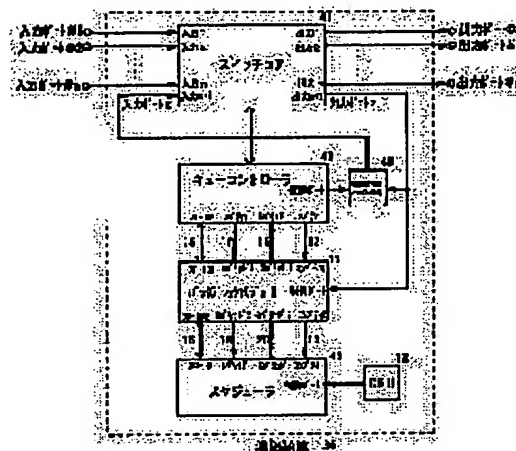
(72)Inventor : YAMASHITA KAZUHISA

## (54) COMMUNICATION EQUIPMENT

## (57)Abstract:

PROBLEM TO BE SOLVED: To actualize selective back pressure on which efficient flow control is based only as to some selected flows.

SOLUTION: A packed or cell inputted from the input port of a switch core 41 is temporarily stored in a queue provided for each flow and then outputted from the output port through a multiplexer. When a trailing device in the communication equipment makes a request to stop a flow, a back pressure module 11 stops the flow by sending a notice showing that 'the queue is empty' to a scheduler 43. At a request to restart the flow which is made by the trailing device, the back pressure module 11 restarts the flow by sending a notice indicating that 'the queue is not empty' to the scheduler 4.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office



F1355

特願平2001-129137号③

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-183886

(P2000-183886A)

(43) 公開日 平成12年6月30日 (2000.6.30)

(51) Int.Cl.<sup>7</sup>

H 0 4 L 12/28

識別記号

F I

H 0 4 L 11/20

タームト\* (参考)

G 5 K 0 3 0

9 A 0 0 1

審査請求 未請求 請求項の数3 O L (全 14 頁)

(21) 出願番号

特願平10-360232

(22) 出願日

平成10年12月18日 (1998.12.18)

(71) 出願人 000002130

住友電気工業株式会社

大阪府大阪市中央区北浜四丁目5番33号

(72) 発明者 山下 和寿

大阪府大阪市此花区島屋一丁目1番3号

住友電気工業株式会社大阪製作所内

(74) 代理人 100078813

弁理士 上代 哲司 (外2名)

Fターム (参考) 5K030 HA08 HA10 HB01 HB02 HB28

HB29 KA02 MB01 MB09

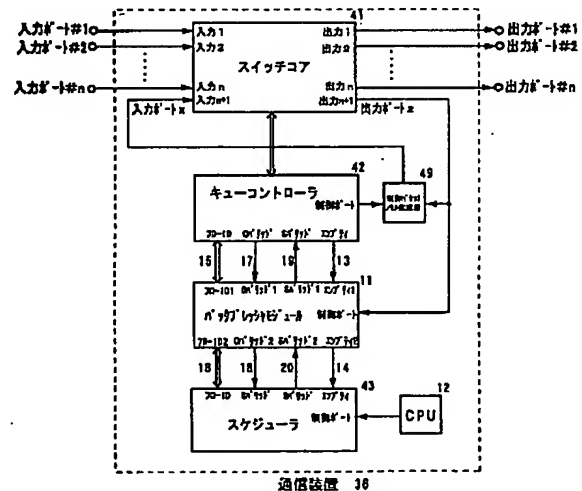
9A001 CC07 JJ18 KK56

(54) 【発明の名称】 通信装置

(57) 【要約】

【課題】 一部の選択されたフローに関してのみバックプレッシャに基づくフロー制御を効率良く行う選択的バックプレッシャを実現する。

【解決手段】 スイッチコア41の入力ポートから入力されたパケット又はセルは、フロー毎に設けたキューに一旦記憶された後、マルチプレクサを経て出力ポートから出力される。本通信装置の後段の装置からフローの停止要求があると、バックプレッシャモジュール11は、「キューが空になった」旨の通知をスケジューラ43に送ることによってフローを止める。後段の装置からフローの再開要求があると、バックプレッシャモジュール11は、「キューが空でなくなった」旨の通知をスケジューラ43に送ることによってフローを再開する。



## 【特許請求の範囲】

【請求項 1】 バケット又はセルを入力し、複数のフローそれぞれに対応して設けてあるキューに一旦記憶した後に出力する通信装置において、前記キューの実際の状態を通知するキューコントローラと、前記キューの擬似的な状態を通知するバックプレッシャモジュールであって、出力先から前記フローのバケット又はセルの出力を停止する要求があれば擬似的な状態を空とし、出力先から前記フローのバケット又はセルの出力を再開する要求があれば擬似的な状態を空でないとする、バックプレッシャモジュールと、前記キューコントローラ及び前記バックプレッシャモジュールから通知を受け、前記実際の状態又は前記擬似的な状態が空の場合はバケット又はセルをキューから取り出さず、前記実際の状態及び前記擬似的な状態が空でない場合はバケット又はセルをキューから取り出す制御を行うスケジューラと、を有することを特徴とする通信装置。

【請求項 2】 バケット又はセルを入力し、複数のフローそれぞれに対応して設けてあるキューに一旦記憶した後に出力する通信装置において、前記キューの状態をバックプレッシャモジュールへ通知するキューコントローラと、出力先から前記フローのバケット又はセルの出力を停止する要求があれば、キューが空の状態をスケジューラに通知し、出力先から前記フローのバケット又はセルの出力を再開する要求があれば、前記キューコントローラからの前記キューの状態をスケジューラへ通知するバックプレッシャモジュールと、前記バックプレッシャモジュールから受けた通知によりキューの状態が空の場合はバケット又はセルをキューから取り出さず、前記バックプレッシャモジュールから受けたキューの状態が空でない場合はバケット又はセルをキューから取り出す制御を行うスケジューラと、を有することを特徴とする通信装置。

【請求項 3】 請求項 1 又は 2 に記載の通信装置において、出力先から前記フローのバケット又はセルの出力を停止又は再開する要求が、バケット又はセルによって送付されるものであり、このバケット又はセルをバックプレッシャモジュールが受け取って前記要求を判断することを特徴とする通信装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、デジタルデータを伝送する際のバケットや ATM (Asynchronous Transfer Mode: 非同期転送モード) セルを中継する通信装置に関する。

## 【0002】

【従来の技術】 コンピュータやデータ端末などを通信装置で接続してネットワークを形成し、各コンピュータ間でデータを伝送することが行われている。図 3 にその構成を例示する。コンピュータ 31～34 は、直接的には

通信装置 35～37 に接続されている。通信は双方向に行われるので、接続には少なくとも一対の伝送路が必要である。接続は、複数対の伝送路を用いても良く、ネットワーク全体として迂回路やループがあっても良い。後の説明の都合上、図 3 の通信装置では、左側に入力ポートを、右側に出力ポートを記載している。通信装置 35～37 は、互いに接続されており、コンピュータからの通信を他のコンピュータ又は通信装置に伝送する。この様に通信装置で中継することにより、コンピュータ間での通信を可能にする。通信は、データをバケットに区切り、バケットを単位として伝送することで行う。ATM 技術を使う場合は、バケットをさらに細分化したセルを単位として伝送する。本発明に関する通信装置は、バケットを伝送する場合でもセルを伝送する場合でも適用可能である。バケットとセルと特に区別して説明する必要がないので、以下では、「バケット又はセル」を「バケット／セル」と表記する。

【0003】 図 3 のネットワークでは、通信装置を介してコンピュータとコンピュータの間で通信が行われる。この通信の流れをフローと呼ぶ。通信の対象となるデータには、リアルタイムの動画や音声の様に、通信の遅延や伝送レートが一定以下に劣化すると困るものがある。また、一方、ファイル転送のようになるべく高い伝送レートで送りたいが、遅延や伝送レートの劣化があっても差し支えないものがある。前者は、QoS (Quality of Service: サービス品質) が一定以上に保たれる必要がある。後者は、QoS を一定以上に保つ必要がない。一方、ネットワークの使用効率はなるべく高めたい。すなわち、通信できる最大の伝送レート (伝送帯域) で、通信したい。

【0004】 そこで、通信装置においては、まず QoS を一定以上に保つべき通信に伝送レートを割り当てた後、残る伝送帯域については QoS を保たなくて良い通信に伝送レートを割り当てる。この様に、通信の対象となるデータの性質に合わせて伝送レートを割り当てることで、ネットワークの使用効率を高めることができる。

【0005】 さて、途中でバケット／セルが廃棄されることなく装置間で通信が行われるためには、後段の装置が受け取ることができる量に合わせて前段の装置が送信を行わなければならない。これをフロー制御と呼ぶ。具体的には、後段の装置がバケットを受け取れるか否かをその前段の装置に通知することで行う。例えば、後段の装置内に設けてあるバケット／セルを受け取るバッファ (一時的な記憶領域) に溜まっているバケット／セルの量を監視し、所定の閾値以上になれば、後段の装置から前段の装置へ送信を停止するよう通知する。逆に、バッファに溜まっているバケットの量が、別の所定の閾値以下になれば、後段の装置から前段の装置へ送信を再開するよう通知する。この様な、後段の状態による前段の送信の制御をバックプレッシャと呼ぶ。

【0006】ところで、通信装置の間をつなぐ伝送路には、コンピュータ間の通信が複数含まれる。フローに必要なQoSは前述した様に種々に異なりうる。すなわち、QoSの異なる複数のフローが同じ通信装置の経路を通る。また、バックプレッシャの生じ方は個々のフロー毎に異なる。従って、フロー毎にバックプレッシャを制御できる必要がある。これを選択的バックプレッシャと呼ぶ。

【0007】従来技術により選択的バックプレッシャを行う通信装置36の構成を図4に示す。通信装置36は、n本の入力ポートとn本の出力ポートを備えており、n個の別の装置と通信を行うことができる。

【0008】41はスイッチコアであり、n+1本の入力ポートとn+1本の出力ポートを備える。n+1番目の入力ポート（入力ポートx）と出力ポート（出力ポートz）は、この通信装置36自身が、別の装置と通信するために用いる。例えば、選択的バックプレッシャに基づくフロー制御のためには、この通信装置と別の装置の間での連携が必要である。この入力ポートxと出力ポートzを使って制御用パケット／セルを通信することで、この連携を取る。

【0009】スイッチコア41に入ってから出るまでのパケットの流れを図5に模式的に示す。入力されたパケット／セルは、クラシフィケーション51～53によって、どのフローに対応するものかが特定され、どの出力ポートから出力すべきかが定められる。クラシフィケーション51～53で特定されたフローは、本通信装置内ではフローID（Identifier）を付して表すことにする。パケット／セルは、フロー毎に設けたキュー（待ち行列）55～59に一旦蓄えられる。キューは、FIFO（First In First Out：先に入れたものから先に取り出される）形式で動作する記憶手段で実施される。スイッチマトリックス54は、フローIDに基づいてパケット／セルをそれぞれのキューに割り振って書き込む。マルチプレクサ61は、フロー毎に設けたキュー55～57から、逐次的にパケット／セルを取り出すことで時分割多重化して出力する機能を有する。マルチプレクサとフロー毎のキューは、出力ポートそれぞれに対して用意されている（64～66）。

【0010】図4のキューコントローラ42は、スイッチマトリックス54を経てキューにパケット／セルを入れる操作や、マルチプレクサがキューからパケット／セルを取り出す操作を指示する。すなわち、キューコントローラ42からは、どのキューに入れるべきかが伝えられ、これに基づいてパケット／セルがキューに入れられる。また、キューコントローラ42は、出力すべきパケット／セルをスイッチコア41に伝える。これらに基づいて、スイッチコア41内部のマルチプレクサ61～66は、パケット／セルをキュー55～59から取り出して所定の出力ポートから出力する。

【0011】図5で模式的に示した様に、パケット／セルの全体をキューに入れる構成も可能である。しかし、キューの記憶場所の節約やキューの処理を容易にするため、パケット／セルの順序を記憶するキューとは別に、パケット／セルを記憶する共用メモリを設けるのでも良い。この場合の構成を図6に示す。図6において、共用メモリ71は、共用メモリコントローラ72によって書き込みと読み出しが制御されている。共用メモリ71のどこにパケット／セルを記憶したかは、共用メモリコントローラ72が生成するパケット識別子で表すこととする。キューコントローラ42の内部には、フロー毎に対応したキューを設けておき、このキューにパケット識別子を記憶する。パケット識別子の大きさ（バイト数）は、パケット／セル全体に比べて小さいので、パケット／セル全体をキューに記憶する場合に比べて、キューの記憶や操作が容易である。

【0012】図6において、クラシフィケーション51～53によってフローが特定されたパケット／セルがスイッチマトリックス54に入力されたら、スイッチマトリックス54はそのフローIDをキューコントローラに通知する。一方では、スイッチマトリックス54は、共用メモリコントローラ72を制御してそのパケット／セルを共用メモリ71に書き込む。そして、共用メモリコントローラ72は、パケット識別子をキューコントローラ42に通知する。キューコントローラ42は、フローIDとパケット識別子を受け取って、フローIDで示されるキューにパケット識別子を書き込む。キューからパケット識別子を取り出されるときは、そのパケット識別子とキューに付されたフローIDがキューコントローラ42から共用メモリコントローラ72に通知され、共用メモリ71から該当するパケット／セルが取り出される。取り出されたパケット／セルはフローIDに基づいて対応する出力バッファ73～75に一旦書き込まれた後、出力される。

【0013】以下では、説明の容易化のために、パケット／セルを図5のキュー55～59に記憶する場合ではなく、図6の共有メモリ71に記憶する場合を代表して説明する。

【0014】図4のキューコントローラ42とスケジューラ43の間の通信は、双方向の信号線であるフローID信号44、この信号線を制御するQバリッド信号45、Sバリッド信号46、及び、エンブティ信号47によって、行われる。キューコントローラ42からスケジューラへフローIDを伝えるときは、キューコントローラ42がフローIDを双方向の信号線である44に出力し、Qバリッド信号45をアサートする。逆に、スケジューラからキューコントローラ42にフローIDを伝えるときは、キューコントローラ42がフローIDを双方向の信号線である44に出力し、Sバリッド信号46をアサートする。

【0015】キューコントローラ42は、キューに入っているバケット識別子の状態の変化をスケジューラ43に伝える。キューが空（キューに記憶されているバケット識別子がない状態）のときに、新たなバケット識別子がキューに入れられたなら、キューコントローラ42はスケジューラ43にその旨を通知する。具体的には、フローIDを双方向の信号線である44に出力し、Qバリッド信号45をアサートする。

【0016】スケジューラ43はキューコントローラ42に、出力すべきフローに対応したフローIDを伝える。具体的には、フローIDを双方向の信号線である44に出力し、Sバリッド信号46をアサートする。キューコントローラ42は、キューからバケット識別子を取り出して、フローIDと共に共用メモリコントローラ72に送り、そのバケット／セルを共用メモリ71から取り出して出力バッファ73を経て出力させる。キューからバケット識別子を取り出すことでキューが空になった場合は、キューコントローラ42はスケジューラ43にその旨を通知する。具体的には、キューコントローラ42は、エンブティ信号47をアサートする。

【0017】本例では、一組のキューコントローラ42とスケジューラ43が、全ての出力ポートに対応する、キューを管理する場合を示している。キューそれぞれにバケット識別子を入れたり取り出したりする指示は、決められた周期で逐次的に行っている。すなわち、キューコントローラ42とスケジューラ43は、時分割で各ポートに対するサービスを行っている。共用メモリ71から出力バッファ73～75へのバケット／セルの転送は、出力バッファ73～75からバケット／セルが出力ポートに送出されるよりも速く行われる。つまり、共用メモリ71から出力バッファ73～75への転送は時分割で行われるが、出力バッファ73～75からバケット／セルが送出されるのは伝送路に対して定められた速度で連続して行われる。なお、出力ポートからバケット／セルを出力する動作は、互いに独立しているので、キューコントローラ及びスケジューラは出力ポート毎に用意しても良い。

【0018】スケジューラ43の機能について、より詳しく説明する。スケジューラ43は、フローIDと出力ポートとの対応、出力ポートの有する伝送帯域から各フローにどれだけの伝送帯域を割り当てるか、を管理する機能を有している。スケジューラ43は、キューコントローラ42からフローIDを受け取ることで「キューが空でなくなった」ことを知ると、そのフローIDで示されるキューに対する指示（サービス）を、スケジューラ43の処理対象に加える。サービスの対象となっているキューに関しては、伝送帯域を割り当てる所定の方針（ポリシー）に従って、バケット／セルを取り出すタイミングをスケジューラ43は決める。このタイミングになれば、対応するフローIDをスケジューラ43はキュー

ーコントローラ42に伝え、これによりスイッチコア41ではキューからバケット／セルが取り出される。この際、もし、キューコントローラ42からのエンブティ信号47がアサートされたなら、「キューが空になった」ことをスケジューラ43が知り、このキューはサービスの対象から除外される。

【0019】伝送帯域を割り当てるポリシーの代表的なものとして、バケットの場合はDDR (Deficit Round Robin)が、セルの場合はWRR (Weighted Round Robin)がある。これらは、各フローに重みを設定し、その重みの割合に従って、出力ポートの伝送帯域をフローの間で分け合うというものである。ラウンドロビンによる伝送帯域の割り当ては、例えば、サービスの対象となっている全てのキューを循環的に検査し、それぞれのキューからはフローの重みに比例した数だけバケット／セルを出力させることで実現する。

【0020】図4のCPU (Central Processing Unit: マイクロコンピュータ等により実施される中央処理装置) 48は、通信装置36全体を管理する。特に、バックプレッシャに関しては、ポートzから受けた制御用バケット／セルを解釈してフローを制御する機能を有する。すなわち、出力先からバックプレッシャを受けて出力を停止するフローは、スケジューラ43に設けてある制御ポートを通して、CPU 48がそのフローの重みを0にする。出力先のバックプレッシャが解除されて出力を再開するフローについては、CPU 48がそのフローの重みを通常の値に戻す。この様に、スケジューラ43内部にあるフローの重みをCPU 48が書き換えることで、バックプレッシャによるフローの停止・再開を制御する。

【0021】図4の制御バケット／セル生成部49は、通信装置36が他の機器に出力する制御用バケット／セルを生成する。例えば、通信装置36内のキューに記憶されているバケット／セルの量に応じてバックプレッシャによるフローの停止・再開を、本通信装置36の入力先へ伝えるための制御用バケット／セルを生成する。制御バケット／セル生成部49を別に設けずに、CPU 48がこの働きを兼ね備える構成も、もちろん可能である。

【0022】

【発明が解決しようとする課題】かかる構成の通信装置では、各フローの停止・再開によって選択的バックプレッシャを実現しようとする、後段の通信装置には、大きなバッファ（例えば、キューや共用メモリ）が必要である。以下に、より詳述する。

【0023】もし、後段の通信装置からのフローの停止要求が即応されず、バケット／セルの出力が続くと、後段の通信装置はそれだけ余分なバケット／セルを一時的に受け取らなければならない。さもないと、バッファのオーバーフローのため、バケット／セルの廃棄が生じ

る。これを避けるためには、バッファの容量を大きくする必要がある。

【0024】一般に、フロー*i*の伝送レートを $R_i$ とし、後段の通信装置がフロー停止又は再開を要求する制御用パケット／セルを送信して、本通信装置から実際にフローが停止又は再開するまでのラウンドトリップ時間を $T_{rtt}$ とする。後段の通信装置のバッファがバックプレッシャすべきと判断する閾値（停止閾値）を越えた後\*

$$R_i \times (T_{rtt} + T_{procN} + T_{procS})$$

だけのバッファが必要である。

【0025】逆に、後段の通信装置からのフローの再開要求に即応できずにパケット／セルを停止し続けると、後段の通信装置のバッファにあったパケット／セルはさらに次段の通信装置に送られて一時的に減ってしまう。もし、この後段の通信装置のバッファが空になったなら、フローが一時的に止まってしまう。すなわち、アンダーランが生じて、パケット／セルの流れのQoSが保\*

$$R_i \times (T_{rtt} + T_{procN} + T_{procS})$$

だけのバッファが必要である。

【0027】選択的バックプレッシャは、従来技術においては、スケジューラ43の帯域割り当ての設定を帯域0（バックプレッシャによるフロー停止時）、本来の帯域（フローを停止していない時）のいずれかに変更することで実現している。重みを変更するには、スケジューラ43から変更前の重みを読み取って記憶した後に、変更後の重みを書き込む等の処理を行う時間や、重みを書き換えるタイミングと重みを検査で参照するタイミングとが重ならない様にタイミングを制御する時間等が必要である。すなわち、各フローの重み付けの設定変更には時間がかかる。つまり、 $T_{procS}$ と $T_{procN}$ が大きく、従来技術による通信装置は後段の通信装置からのバックプレッシャに即応することができない。そのため、従来技術による通信装置では、フローの停止・再開の応答性が悪く、次段の通信装置には大きなバッファ容量を必要としていた。

【0028】また、ラウンドロビンによる伝送帯域の割り当てのため、前述の様にサービスの対象となっている全てのキューを循環的に検査すると、重みが0のフローに対しても検査が行われることになり、その検査時間が無駄である。重み0のフローが多くなると、実際には出力しないフローの検査にスケジューラ43の時間が使われてしまい、本来出力すべきフローを十分には処理できなくなってしまう。

【0029】本発明は、これらの点を解決するために創作されたものであり、選択的バックプレッシャによるフローの停止・再開を効率良く行うことを目的としている。

【0030】

【課題を解決するための手段】かかる目的に対して、請求項1の通信装置は、パケット又はセルを入力し、複数

\*フロー停止を要求する制御用パケット／セルを生成して本通信装置に送信するまでに必要な処理時間を $T_{procN}$ とし、本通信装置でフロー停止を要求する制御用パケット／セルを受信してから、実際にフローを止めるまでに必要な処理時間を $T_{procS}$ とする。すると、後段の装置においてはフロー*i*のキューをオーバーフローさせないためには、停止閾値の上に、

(1)

10※でないことになる。これを避けるためには、後段の通信装置のバッファの容量を大きくし、パケット／セルが常に多めに残っている様にしなければならない。

【0026】一般に、本通信装置でフロー再開を要求する制御用パケット／セルを受信してから、実際にフローを再開するまでに必要な処理時間を $T_{procR}$ とする。すると、次段の装置においてはフロー*i*を途切れさせないためには、再開閾値の下に、

(2)

のフローそれぞれに対応して設けてあるキューに一旦記憶した後に出力する通信装置において、前記キューの実際の状態を通知するキューコントローラと、前記キューの擬似的な状態を通知するバックプレッシャモジュールであって、出力先から前記フローのパケット又はセルの出力を停止する要求があれば擬似的な状態を空とし、出力先から前記フローのパケット又はセルの出力を再開する要求があれば擬似的な状態を空でないとする、バックプレッシャモジュールと、前記キューコントローラ及び前記バックプレッシャモジュールから通知を受け、前記実際の状態又は前記擬似的な状態が空の場合はパケット又はセルをキューから取り出さず、前記実際の状態及び前記擬似的な状態が空でない場合はパケット又はセルをキューから取り出す制御を行うスケジューラと、を有することを特徴とする。

【0031】かかる様に構成されているので、スケジューラはキューコントローラからの通知と同様にバックプレッシャモジュールからの通知にも従う。これにより、キューコントローラの通知による場合と同様のやり方で、バックプレッシャモジュールの通知によるパケット又はセルの出力の停止・再開が制御できる。

40 【0032】請求項2の通信装置は、パケット又はセルを入力し、複数のフローそれぞれに対応して設けてあるキューに一旦記憶した後に出力する通信装置において、前記キューの状態をバックプレッシャモジュールへ通知するキューコントローラと、出力先から前記フローのパケット又はセルの出力を停止する要求があれば、キューが空の状態をスケジューラに通知し、出力先から前記フローのパケット又はセルの出力を再開する要求があれば、前記キューコントローラからの前記キューの状態をスケジューラへ通知するバックプレッシャモジュールと、前記バックプレッシャモジュールから受けた通知に

よりキューの状態が空の場合はバケット又はセルをキューから取り出さず、前記バックブレッシャモジュールから受けたキューの状態が空でない場合はバケット又はセルをキューから取り出す制御を行うスケジューラと、を有することを特徴とする。

【0033】かかる様に構成されているので、スケジューラはキューコントローラの発する通知と共にバックブレッシャモジュールの発する通知を、バックブレッシャモジュールから受け取る。これにより、キューコントローラの通知による場合と同様のやり方で、バックブレッシャモジュールの通知によるバケット又はセルの出力の停止・再開が制御できる。さらに、キューコントローラから通知を受ける信号線とは別にバックブレッシャモジュールから通知を受ける信号線をスケジューラに設ける必要がない。

【0034】請求項3の通信装置は、請求項1又は2に記載の通信装置において、出力先から前記フローのバケット又はセルの出力を停止又は再開する要求が、バケット又はセルによって送付されるものであり、このバケット又はセルをバックブレッシャモジュールが受け取って前記要求を判断することを特徴とする。

【0035】かかる様に構成されているので、バックブレッシャモジュールは出力先からフローの停止・再開要求をバケットで受け取ってすみやかに判断することができる。

【0036】

【発明の実施の形態】本発明の実施例を図1に示す。この実施例は、従来のキューコントローラ及びスケジューラが有していた信号線を利用して、選択的バックブレッシャによるフローの停止及び再開を行っている。そのため、従来のキューコントローラ及びスケジューラを変更する必要がなく、装置の設計が容易である。

【0037】図1において、36以上の符号を付した部分は、図4における同符号の部分と同じ機能構成である。すなわち、41はスイッチコアであり、入力ポートからバケット／セルを入力し、フロー毎に設けたキューに蓄えた後、マルチプレクサによる時分割多重を経て、出力ポートから出力する。図5の様にスイッチコア41内部にバケット／セル全体を記憶するキューを設ける他、図6に示した様に、バケット／セルを示すバケット識別子を作成してキューコントローラ42内に設けたキューにバケット識別子を記憶することでも本発明は実施できる。キューコントローラ42とスケジューラ43をつなぐ信号線には途中にバックブレッシャモジュール11を設けてある。

【0038】図6のキューコントローラ42は、内部に設けたフロー毎のキューにバケット識別子を記憶し、スイッチコア41（具体的には、図6の共用メモリコントローラ72）に対してバケット／セルの出力を指示する。さらに、キューコントローラ42は、キューの状態

を図1のバックブレッシャモジュール11を経てスケジューラ43へ伝える。スケジューラ43は、各フローに対して設定した重み付けに従って出力ポートの伝送帯域を各フローに割り当て、キューからバケット／セルを取り出すタイミングをバックブレッシャモジュール11を経てキューコントローラ42へ伝える。制御バケット／セル生成部49は、通信装置36から他の機器に伝送する制御のためのバケット／セルを生成する。

【0039】選択的バックブレッシャによるフローの停止・再開を行わない時は、バックブレッシャモジュール11は、キューコントローラ42とスケジューラ43の間の通知を素通しする（信号をそのまま通す場合の他に、通知を一旦受け取ってから送る場合も含む）。選択的バックブレッシャを行うために、バックブレッシャモジュール11は、出力ポート2から制御用バケット／セルを受け取り、フローを停止・再開させるべきキューを判断する。

【0040】従来技術で述べた様に、キューコントローラ42は、スケジューラ43からフローIDを受け取ると、キューコントローラ42内部の対応するキューからバケット識別子を取り出し、スイッチコア41に指示してバケット／セルを出力させる。この時、キューが空でなければキューコントローラ42はエンプティ信号13をアサートしないが、キューが空になればエンプティ信号13をアサートする。

【0041】バックブレッシャモジュール11は、受け取った制御用バケット／セルに基づいてフローを停止すべきと判断したときは、キューコントローラ42がスケジューラ43からフローIDを次に受け取るタイミングに合わせて、エンプティ信号14をアサートする。すなわち、実際のキューが空でなくても（エンプティ信号13がアサートされなくても）、バックブレッシャモジュール11は「キューは空になった」という通知をスケジューラ43に送る。すると、スケジューラ43は、そのフローをサービス対象から除外するので、以降はそのフローの出力が行われなくなる。

【0042】キューコントローラ42とスケジューラ43の間の通信は、双方向の信号線であるフローID信号15と16、Qバリッド信号17と18、Sバリッド信号19と20を用いて、時分割で行われる。この通信は、バケット／セルのタイミングに同期して生じるので、時間的な隙間がある。

【0043】バックブレッシャモジュール11は、受け取った制御用バケット／セルに基づいてフローを再開すべきと判断したときは、この隙間を使って、スケジューラ43に、フローIDを通知する。すなわち、キューコントローラ42からの双方向の信号線である15にはフローIDが示されず、Qバリッド信号17がアサートされていないにも拘わらず、バックブレッシャモジュール11は、フローIDを双方向の信号線である16に示



し、Qバリッド信号18をアサートする。

【0044】すると、バックブレッシャモジュール11は、「キューは空でなくなった」と判断してそのフローをサービス対象に加えるので、以降はそのフローの出力が行われる。

【0045】この様に、バックブレッシャモジュール11は、選択的バックブレッシャによるフローの停止・再開の操作を、スケジューラ43への「キューは空になった」・「キューは空でなくなった」という擬似的な通知によって実現する。

【0046】なお、選択的バックブレッシャによりフローの停止・再開の操作を行う以外の時は、バックブレッシャモジュール11は、スケジューラ43からの通知をそのままキューコントローラ42に送り、キューコントローラ42からの通知はそのままスケジューラ43に送る。すなわち、通知を素通しするので、従来技術と同様に「キューは空になった」・「キューは空でなくなった」という実際の通知に応じてフローの制御は行われる。

【0047】以上を整理して説明する。バックブレッシャモジュール11からスケジューラ43に通知される擬似的な状態が「キューは空である」又は、キューコントローラ42からスケジューラ43に通知される実際の状態が「キューは空である」ならば、スケジューラ43はそのフローを制御対象から除外しているので、そのフローのバケット／セルは出力されない。逆に、バックブレッシャモジュール11からスケジューラ43に通知される擬似的な状態が「キューは空でない」かつ、キューコントローラ42からスケジューラ43に通知される実際の状態が「キューは空でない」ならば、スケジューラ43はそのフローを制御対象に含めているので、そのフローのバケット／セルは出力される。

【0048】ここで、「キューは空である」あるいは「キューは空でない」という状態は、「キューは空になった」あるいは「キューは空でなくなった」という通知を受けることによってスケジューラは把握している。キューコントローラ42からのこの通知は、バックブレッシャモジュール11を経てスケジューラ43に伝えられる。フローの停止・再開のためのこの通知はバックブレッシャモジュール11からスケジューラ43に伝えられる。

【0049】なお、キューコントローラ42内のキューが空になった直後に、フローの再開を要求する制御用バケット／セルをバックブレッシャモジュール11が受け取る場合がある。この場合は、キューの実際の状態に鑑みて、「キューは空でなくなった」という通知をバックブレッシャモジュール11は発行しなくても良い。あるいは、あえてバックブレッシャモジュール11からスケジューラ43に「キューは空でなくなった」と通知しておき、スケジューラ43がそのキューからのバケット識

別子の取り出しをキューコントローラ42に指示したときに、キューコントローラ42がこの指示を無視する一方「キューは空になった」ことをスケジューラ43へ通知し直すのも良い。

【0050】12は装置全体を管理するCPUである。本実施例の場合は、初期設定を除いて、バックブレッシャによるフローの停止・再開処理を装置の稼働時にCPU12が行う必要はない。CPU12の処理すべき事項が図4に比べて少ないので、このCPUを他の用途にも用いることが可能である。また、制御バケット／セル生成部49が行う制御用バケット／セルの生成処理をこのCPU12に行わせる実施例も可能である。

【0051】以上の説明において、キューコントローラ42、バックブレッシャモジュール11、スケジューラ43間のフローID信号15と16は、双方向の信号線になっている。これは、キューコントローラ42からバックブレッシャモジュール11を経てスケジューラ43へフローIDが通知される期間と、逆にスケジューラ43からバックブレッシャモジュール11を経てキューコントローラ42へフローIDが通知される期間が明瞭に分離されている場合の実施例である。もちろん、これ以外に、フローIDの信号を方向別に分け、両方向の通知を同時に行いうる様にしても良い。

【0052】他の実施例として、出力ポートzからの制御バケット／セルを一旦CPUで解釈し、CPUが動作をバックブレッシャモジュールに通知する構成もある。また、図1の実施例ではバックブレッシャモジュール11がキューコントローラ42とスケジューラ43間の信号すべてを分断している様に見えるが、バックブレッシャモジュール11が変更しなくて良い信号は、キューコントローラ42とスケジューラ43とを直接接続しても良い。例えば、Sバリッド信号をバックブレッシャモジュール11は変更しない。これらを考慮した実施例を図2に示す。

【0053】図2において、バックブレッシャモジュール21はSバリッド信号46を参照のみ行う。CPU22は、出力ポートzからの制御バケット／セルを受け取って、フローの停止・再開をすべきフローIDをバックブレッシャモジュール21に通知する。スケジューラ43に対しては、CPU22は図1のCPU12と同様の装置起動時の初期設定を行う。

【0054】また、出力ポートzからの制御バケット／セルにフローの停止・再開要求が複数含まれている場合は、CPU22からバックブレッシャモジュール21へは、フローを停止する複数のフローIDや、フローを再開する複数のフローIDをリスト形式に整理し、一括して通知する実施例も可能である。すると、CPU22からバックブレッシャモジュール21への通知が効率良く行える。

【0055】また、従来技術で述べた様に、出力ポート

10

20

30

40

50



からバケット／セルを出力する動作は、互いに独立している。従って、出力ポート毎に、キューコントローラとバックブレッシャモジュールとスケジューラを設ける実施例も可能である。

【0056】

【発明の効果】本発明では、フローの停止及び再開を、スケジューラの重み設定を変更することなく、キューの状態を擬似的に変更することで行う。そのため、フロー停止を要求する制御用バケット／セルを受信してから実際にフローを止めるまでに必要な処理時間 $T_{proc}$ 、及び、フロー再開を要求する制御用バケット／セルを受信してから実際にフローを再開するまでに必要な処理時間 $T_{proc}$ を小さくすることができる。つまり、(1)式の $T_{proc}$ 及び(2)式の $T_{proc}$ が小さくなるので、本通信装置の出力先に繋がる後段の装置が用意すべきバッファ量を小さくできる。

【0057】また、ラウンドロビンにより循環的にサービスの対象となるフローを検査する場合には、重みを0に設定したフローを含めることがないので、スケジューラの時間が無駄に使われることがない。

【0058】キューコントローラからバックブレッシャモジュールを経由してスケジューラへの通知を行う構成においては、キューコントローラから通知を受ける信号線とは別にバックブレッシャモジュールから通知を受ける信号線をスケジューラに設ける必要がない。そのため、従来技術で用いていたスケジューラをそのまま使うことができるので、設計が容易となる。

【0059】フローの停止及び再開のバケット／セルを、CPUが介在することなく直接バックブレッシャモジュールが受け取る場合には、CPUの処理が不要となつてより高速に処理できるので、さらに $T_{proc}$ 及び $T_{proc}$ を小さくすることができる。そのため、後段の装置が用意すべきバッファ量をさらに小さくできる。

【図面の簡単な説明】

【図1】本発明による通信装置の実施例を示す図である。

【図2】本発明による通信装置の他の実施例を示す図である。

【図3】通信ネットワークの例を示す図である。

【図4】従来技術による通信装置の構成を示す図である。

【図5】バケットの流れを示す模式図である。

【図6】スイッチコア41の内部構成の例を示す図である。

【符号の説明】

11、21……バックブレッシャモジュール

12、22、48……CPU

13、14、47……エンプティ信号

15、16、44……フローID信号

17、18、45……Qバリッド信号

19、20、46……Sバリッド信号

31、32、33、34……コンピュータ

20 35、36、37……通信装置

41……スイッチコア

42……キューコントローラ

43……スケジューラ

49……制御バケット／セル生成部

51、52、53……クラシフィケーション

54……スイッチマトリックス

55、56、57、58、59……キュー

61、62、63……マルチプレクサ

64、65、66……キューとマルチプレクサの組み合わせ

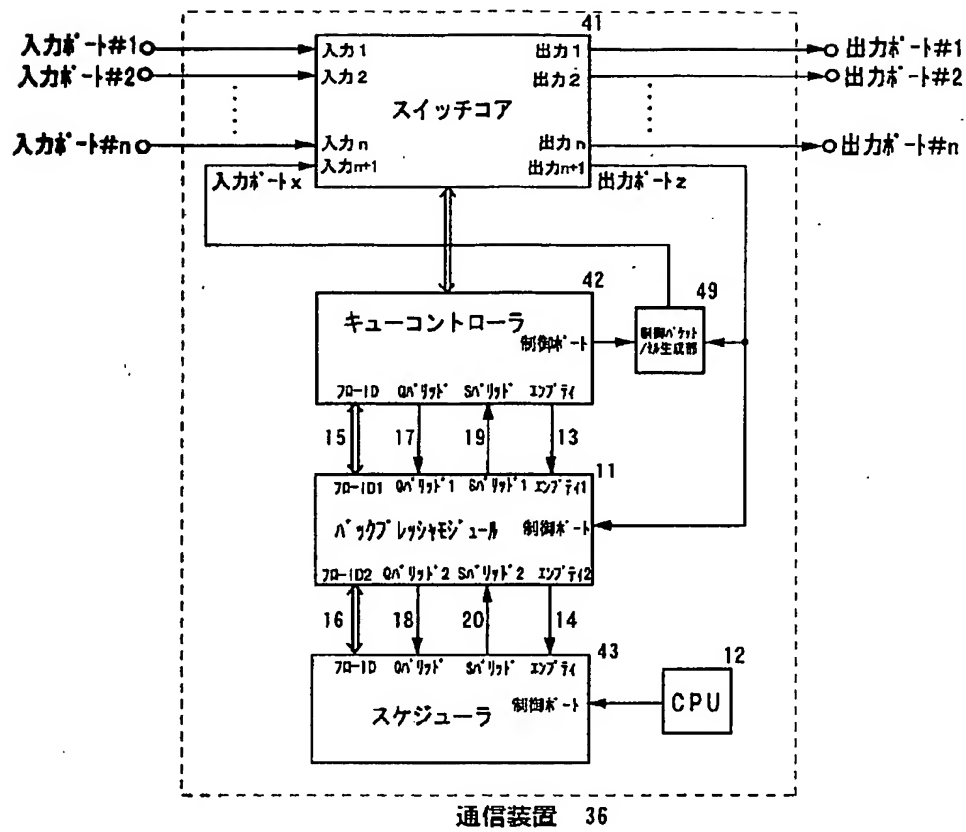
30 わせ

71……共用メモリ

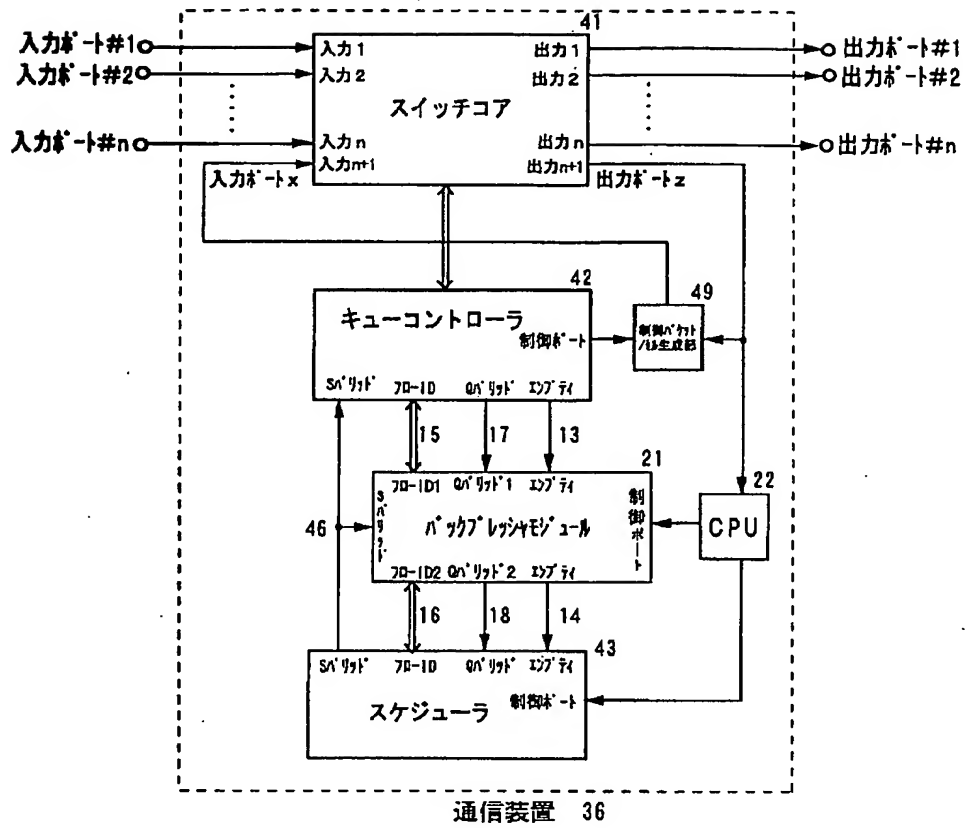
72……共用メモリコントローラ

73、74、75……出力バッファ

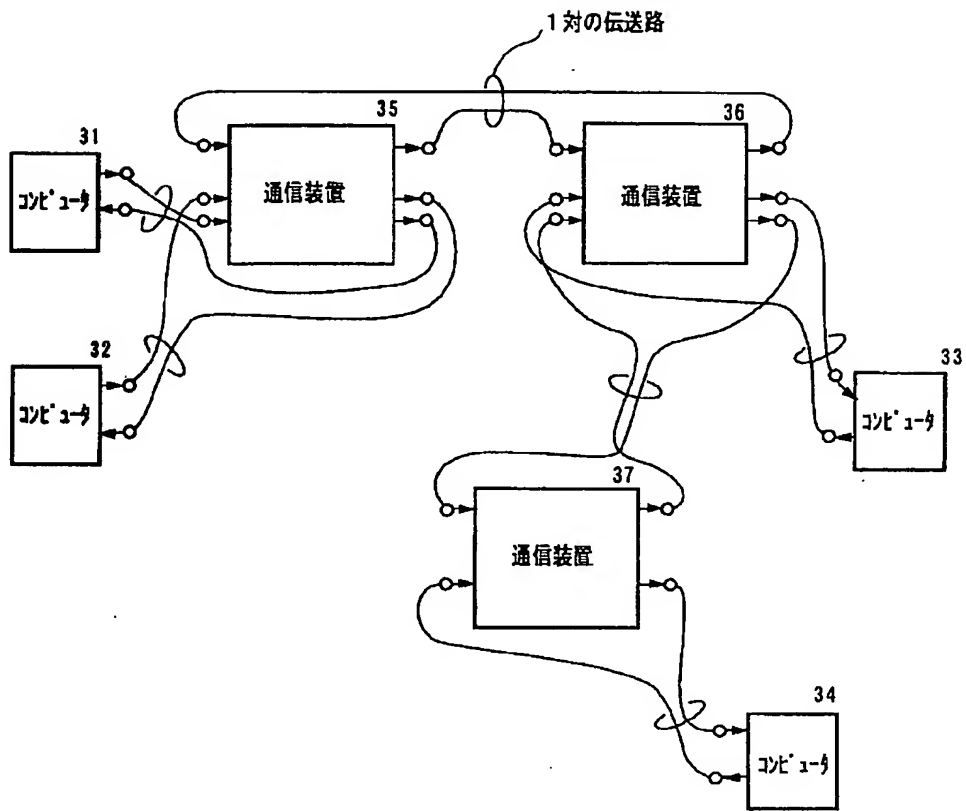
【図1】



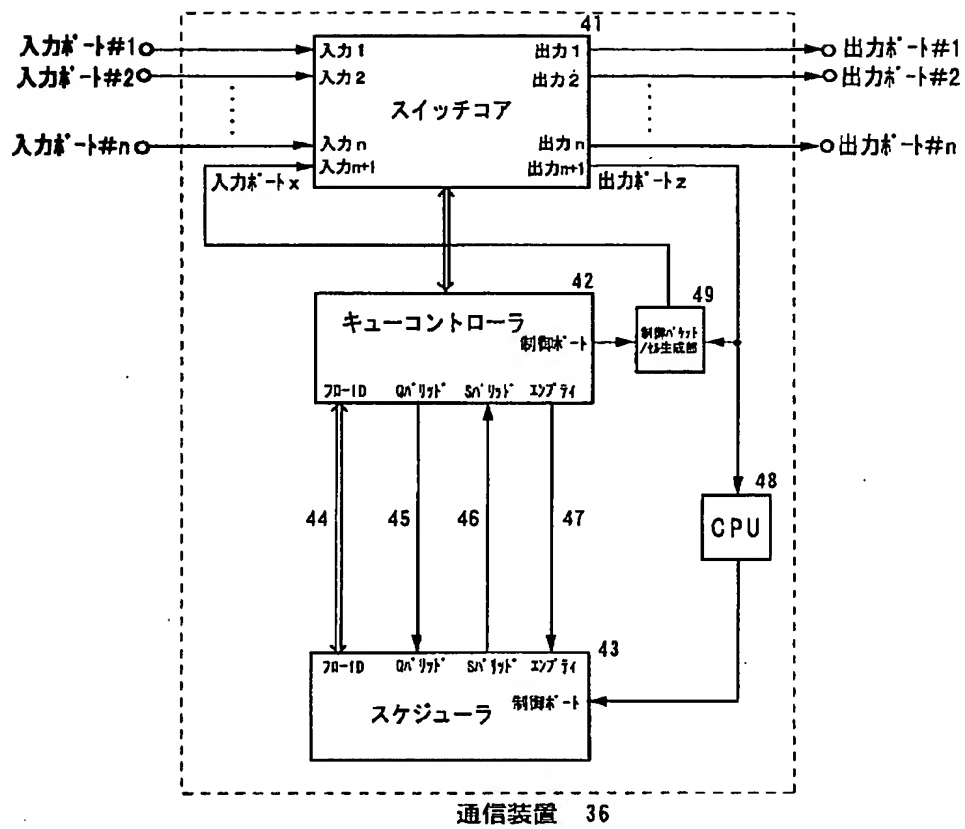
【図2】



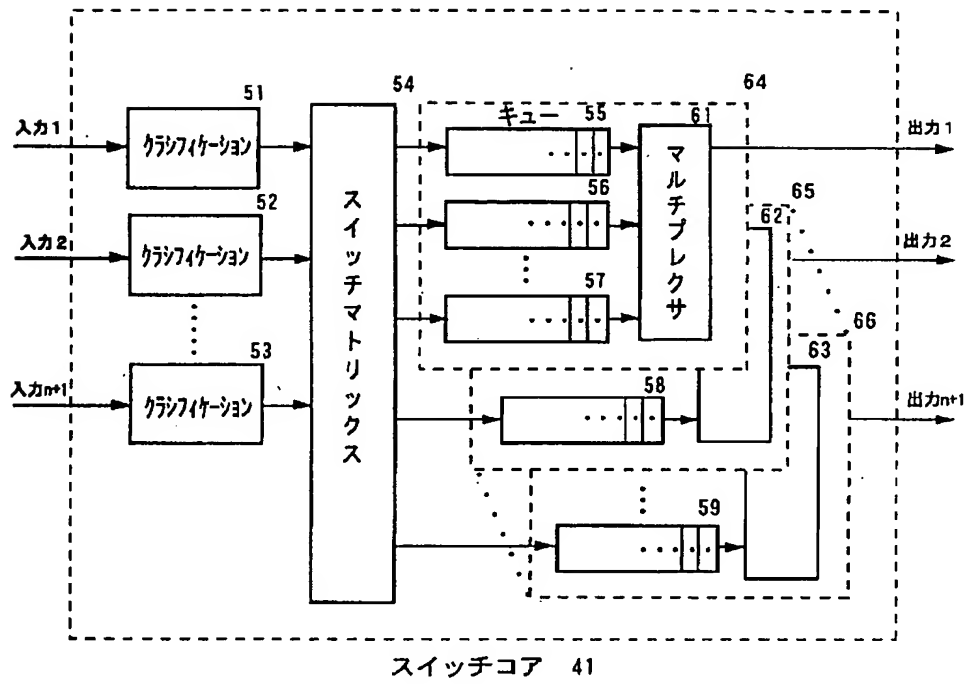
【図3】



【図 4】



【図5】



【図6】

